DOCUMENT RESUME

ED 201 658 TM 810 251

AUTHOR Baranowski, B. Bonnie; Halperin, Silas
TITLE A Comparison of Six Robust Correlation Estimators.
PUB DATE 14 Apr 81
NOTE 35p.; Paper presented at the Annual Meeting of the American Educational Research Association (65th, Los Angeles, CA, April 13-17, 1981).

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Comparative Analysis; *Correlation; Mathematical Formulas; *Statistical Bias
IDENTIFIERS Estimation (Mathematics); *Monte Carlo Studies; *Robustness; Sample Size; Variability

ABSTRACT

A Monte Carlo investigation of six robust correlation estimators was conducted for data from distributions with longer than Gaussian tails: a bisquare coefficient, the Tukey correlation, the standardized sums and differences, a biweight standardized sums and differences, the transformed Spearman's rho and a bivariate trimmed Pearson. Evaluation of the estimators was based on bias and variability as measured by mean square error, and efficiency relative to the Pearson correlation coefficient. Correlations of .9, .3, .6, and .0 for samples of size twenty and forty for different weight tails were used. Two estimators that stood out as not being situation specific and demonstrated robust properties were the averaged bisquare, and Tukey correlations. The transformed Spearman's rho was easy to calculate, and for small correlations was clearly the best of all the estimators considered; however, it degenerated as the correlation increased so that it was among the worst for a correlation of .9. The standardized sums and differences estimator demonstrated robustness and warrants further investigation. The biweight standardized sums and differences estimator did not demonstrate robustness, and the bivariate trimming estimator demonstrated limited usefulness. (Author/RL)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED201658

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

A COMPARISON OF SIX ROBUST CORRELATION ESTIMATORS

Dr. B. Bonnie Baranowski

and

Dr. Silas Halperin

Syracuse University

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

B.B. Baranowski

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

TM 810 251

Presented to the American Educational Research Association

April 14, 1981, Los Angeles, California

ABSTRACT

A Monte Carlo investigation of six robust correlation estimators was conducted for data from distributions with longer than Gaussian tails: a bisquare coefficient, the Tukey correlation, the standardized sums and differences, a biweight standardized sums and differences, the transformed Spearman's rho and a bivariate trimmed Pearson. Evaluation of the estimators was based on bias and variability as measured by mean square error, and efficiency relative to the Pearson correlation coefficient. Correlations of .9, .3, .6, and .0 for samples of size twenty and forty for different weight tails were used.

1. INTRODUCTION

During the past two decades, researchers have been investigating the validity and efficiency of least squares estimators when the assumption of a Gaussian distribution is violated and have clearly demonstrated the inadequacy of such estimators, including the Pearson product-moment correlation coefficient, for distributions with longer than Gaussian tails. A new class of estimators, called resistant and robust, has been proposed to estimate parameters of long tailed distributions (Hampel, 1971; Hogg, 1967; Huber, 1972; Wainer, 1976). A distinction between resistance and robustness is not always clear in the literature and the term robust is often used to describe both properties. Resistance refers to the ability of a summary statistic to remain constant in view of a change in a small part of the data, no matter what part or how substantial the change. The term robustness does not have singular meaning and in the context of this paper refers to the variability of the estimator.

The resistant and robust estimation of parameters involves the transformation of the observations to some other values such as ranks and/or the differential treatment of the observations such as the trimming of extreme observations or the weighting of all observations by some scheme. Resistant and robust location and regression estimators have been studied extensively (Andrews, Bickel,

Hampel, Huber, Rogers and Tukey, 1972; Andrews, 1974; Gross, 1976, 1977; Huber, 1977; Maronna, 1976; Ramsay, 1977), but the investigation of such correlation estimators is limited (Denby and Larson, 1977; Devlin, Gnanadesikan and Kettenring, 1975; Gnanadesikan and Kettenring, 1972; Wainer and Thissen, 1976). Unlike the least squares solution, there is not a unique relationship between correlation and regression for most robust and resistant estimators. The literatures are therefore separate and distinct.

The criteria for the treatment of observations for the estimation of the correlation coefficient must take into account the position/value of an observation within the respective univariate distributions as well as the bivariate distribution. While no one estimator has demonstrated overall superiority, the more promising correlation estimators involve trimming in the univariate context of the X and Y distributions and the bivariate context of the major and minor axes.

This study compares six of the more promising robust correlation estimators. Mean square error and efficiency relative to the Pearson are used to evaluate the estimators.

2. ROBUST CORRELATION ESTIMATION

One of the first papers to give an overview of the problems of multivariate estimation is Gnanadesikan and Kettenring (1972). They base the estimation of correlation on the quarter square identity

$$cov(X,Y) = 1/4\ [\,var(X + Y) - var(X - Y)\,] \qquad 2.1$$

and using the usual standard deviations s(X) and s(Y) obtain the Pearson correlation coefficient,

$$r(P) = cov(X,Y)/s(X)\ s(Y). \qquad 2.2$$

Standardized Sums and Differences - r*(SSD)

Robust correlation estimators can be obtained by using any robust estimators of scale to calculate estimates of the variances and standard deviations and using them in equations 2.1 and 2.2.

A problem with this procedure is that the robust correlation estimates may not lie within the bounds [-1, +1]. The solution is to standardize the univariate distributions first, using robust estimators of location and scale, t* and s*, respectively. Let

$$x = (X-t^*(X))/s^*(x) \text{ and } y = (Y-t^*(Y))/s^*(Y). \qquad 2.3$$

The resulting robust estimator is called the standardized sums and differences, r*(SSD),

$$r^*(SSD) = (s^*(x+y)^2 - s^*(x-y)^2)/(s^*(x+y)^2 + s^*(x-y)^2). \quad 2.4$$

Any robust location and scale estimators can be used but the estimators most often used are 10% symmetrically trimmed means and variances. (In this paper, the trimming percentage refers to the total trimmed, one half of the percent from each end of the distribution.) This estimator performed well in studies by Devlin, et al. and Denby and Larson.

This approach looks first at the univariate distributions independently and then the major (x+y) and minor (x-y) axes of the bivariate space. It is important to note that an observation which is considered a univariate outlier might not be considered a bivariate outlier when we examine the scatter plot of the data and visa versa.

Biweight Standardized Sums and Differences - r*(BSD)

In a study of confidence interval robustness for the correlation coefficient, Chou (1980) developed a biweight measure of correlation having the form of the standardized sums and differences (2.4) but used the biweight location and bar statistics (Mosteller and Tukey, 1977) as the estimators of location and scale. This estimator will be denoted as r*(BSD). In the Chou study it demonstrated robustness in several situations but its properties as a point estimator are unknown.

The biweight statistic is an iterative maximum likelihood estimator of location. Let

$$t^*_{k+1} = \bar{\sum} w_{ik} x_i / \bar{\sum} w_{ik} \qquad 2.5$$

where

$$w_{ik} = (1 - u_i^2)^2 \quad \text{if} \quad u_i^2 \leq 1$$

$$= 0 \quad \text{if} \quad u_i^2 > 1 \qquad 2.6$$

and $$u_i = (X_i - t^*_{k-1}) / cs^*_{k-1} \qquad 2.7$$

and k refers to the kth iteration, c is some constant and s* is a measure of scale.

In this study, c = 9 and the median of the absolute deviations (MAD) was used as the measure of scale.

The corresponding measure of scale is called "bar" for robust variance.

$$\text{bar}(X) = \frac{n \bar{\sum}'(X-t^*)^2 (1-u^2)^4}{[\bar{\sum}'(1-u^2)(1-5u^2)][-1+\bar{\sum}'(1-u^2)(1-5u^2)]} \qquad 2.8$$

where $\bar{\sum}'$ means to sum only those observations where $u^2 \leq 1$.

Tukey Correlation - r*(t)

The Tukey correlation is also based on the quarter square identity where s* is an unbiased estimator of the standard deviation and a robust measure of scale obtained from the order statistics of the observations. Let

$$x'' = X/s^*(x) \qquad \text{and} \qquad y'' = Y/s^*(Y) \qquad 2.9$$

then

$$r^*(t) = 1/4\,[\,(s^*(x'' + y''))^2 - (s^*(x'' - y''))^2\,], \qquad 2.10$$

where

$$s^* = [\,1.77245/(n(n-1))\,] \sum d(i)\, g(i). \qquad 2.11$$

For a sample of size n, the subscript i represents the rank of the observations in ascending order and g(i) represents the distance between two adjacent observations; i.e.,

$$g(i) = X(i + 1) - X(i) \qquad \text{for } i = 1,\ldots,n-1 \qquad 2.12$$

and

$$d(i) = i(n - i). \qquad 2.13$$

The Wainer and Thissen study was inconclusive as to the value of r*(t) and Chou (1979) has shown it to perform well in some situations.

Transformed Spearman's Rho - r*(S)

The transformed Spearman's rho

$$r^*(S) = 2 \sin (3.14159/6)\, r(S) \qquad 2.14$$

is viewed as having robust properties and was included in the study of robust correlation estimation by Wainer and Thissen with mixed results. The ease of calculation and familiarity make it an attractive estimator to use.

Bivariate Trimming - r*(BVT)

Bivariate trimming refers to a general class of robust estimators which trim a specified percent of observations on the major and minor axes. Once a given percent of the sample observations are identified as outliers, they are trimmed and the usual Pearson calculated from the reduced sample. The procedures for identifying the outliers differ. Devlin et al. used an iterative approach and Chou (1979) used the least squares estimates of location and scale to standardize the univariate distributions and then obtained the crossproducts. The largest positive and negative crossproducts were then trimmed.

In this study 10% symmetrically trimmed means and variances were used for standardizing with 20% symmetric trimming of the crossproducts.

Bisquare Coefficient - r*(B')

The bisquare correlation coefficient was introduced by Mosteller and Tukey

$$r^*(B) = \pm 1 / [1 + \text{bar}(Y-bX) / b^2 \text{ bar}(X)]^{1/2} \qquad 2.15$$

with the sign of b, where b is the slope of a bisquare regression fit. The bisquare regression fit results from an iterative weighted least squares algorithm with the weights determined by the relative magnitude of the residuals.

Baranowski (1980) investigated several aspects of this estimator including problems with respect to the preliminary bisquare regression fit. A problem of this estimator is the lack of stability of the estimate given the transposition of X and Y. The use of the least squares start and a constant of 9 was the prefered solution for the regression fit. To solve the problem of transposition, Baranowski proposed an averaged bisquare, r*(B'), obtained by fitting both regression lines, calculating both correlation estimates and taking the average. This averaged bisquare estimate was more efficient than the single bisquare in all situations.

3. DATA and DESIGN

Long tailed distributions were simulated by the generation of contaminated bivariate normal distributions. (This is one of several approaches used in other research and is not intended to imply that this is how long tailed distributions occur in real situations.) Given a bivariate standard normal distribution with a specified correlation, observations were replaced with probability beta with observations from a bivariate normal distribution with mean equal to zero and variance greater than one. The correlation in the bivariate standard normal and the contaminating normal distributions were equal to avoid ambiguity of what is being estimated.

The random normal deviates were generated by the Marsaglia rectangle-wedge method. The random uniform numbers used in the method were generated by the linear congruential method, incorporating the exchange sequence of MacLaren and Marsaglia (Algorithm M; Knuth, 1969) to break the serial correlation between successive numbers. The desired variances and correlations were induced algebraically.

Thirty-two contaminated situations resulting from the combination of four parameters were examined:

1. population correlation: rho = .0, .3, .6 and .9

2. sample size: n = 20 and 40

3. standard deviation of the contaminants: sigma = 2 and 3

4. level of contamination: beta = .10 and .25.

Eight non-contaminated situations were also examined for the combination of sample size and population correlation. These bivariate standard normal situations can be represented as cases where beta equals zero.

For economy not all situations were uniquely generated. Given sample size and population correlation, the same bivariate normal data were used repeatedly across all contaminating situations. This also increases the internal consistency of the comparisons of the estimators across the various contaminated situations. A similar scheme was used by Devlin, et al.

In order to generalize the results of any Monte Carlo study, the variability of the statistics should be small. Prior to the study, it was decided that the maximum standard error of the mean of the correlation estimators should be .01. Since the variability is inversely related to the number of replications, regression sampling (Cochran, 1963), a variance reduction technique, was used in this study to reduce the required number of replications given the desired

precision. The Student's t transformation of the maximum likelihood estimator of the population correlation, given the assumptions of zero mean and equal variances, (Kraemer, 1975) calculated only from the observations from the non-contaminated bivariate distributions was used as the auxiliary variable. Table 1 gives the number of replications used for the various parameter settings.

Insert Table 1 about here

4. RESULTS

To summarize, the following correlation estimators were calculated over forty parameter setting:

1. the Tukey correlation - r*(t),
2. the transformation of Spearman's rho - r*(S),
3. the biweight standardized sums and differences - r*(BSD),
4. the standardized sums and differences - r*(SSD) - using 10% symmetrically trimmed means and variances,
5. bivariate trimming - r*(BVT) - using 10% symmetrically trimmed means and variances for standardizing X and Y and a 20% symmetric trimming of the crossproducts,

6. the averaged bisquare - r*(B') - with a least squares start and constant of 9 for the bisquare regression fit, and

7. the Pearson correlation coefficient - r(P).

Mean square error which takes into account both the variability and bias of an estimator was the principal means of evaluating the robust estimators,

$$MSE = S(r^*)^2 + (\bar{r}^* - rho)^2. \qquad 4.1$$

The MSE'S for the estimators are presented in Table 2 using the two-way analysis by medians (Tukey, 1977). This procedure decomposes the MSE into a grand effect, the effects of the estimators ((r*)) and the effects of the contamination (ct). The values in the subtables are residuals. This type of analysis is especially useful when the residuals are small relative to the other effects since we can focus directly on the effects. Any MSE can be reconstructed by summing the three effects and the residual,

$$MSE = \text{grand effect} + r^* + ct + \text{residual}. \qquad 4.2$$

Insert Table 2 about here

All values of MSE were multiplied by 10000 to make the tables more readable. For each correlation and sample size

there are two tables, one for each of the different contaminating standard deviations. Within each subtable the grand effect is in the upper right, r* effect in the left column and contamination effect in the top row.

The contamination effect is smaller for a standard deviation of two than for three and for ten percent contamination than for twenty-five percent. The residuals for the subtables in Table 2 are not as small as we might like. In some cases, the residuals are larger than the effects indicating that an additive model is not appropriate here and that some sort of interaction might be taking place. No one estimator is best in all situations and the differences between the estimator effects are small, except for an occasional situation where one estimator stands out as best or worst, such as the large negative effects for a correlation of .0.

Insert Table 3 about here

Table 3 contains the grand effects of MSE. MSE varies with the level and extent of contamination as well as sample size and population correlation. To better understand the MSE for the various estimators, bias was examined. Table 4 contains the means, adjusted by the regression sampling procedures and multiplied by 1000, for the Pearson as well as the six robust estimators. Contamination and sample size do not appear to affect bias and since the residuals are

generally close to zero, there does not appear to be any interaction. Since the maximum standard error of the mean is .01 for all estimators, a difference of less than 20 between estimators could be considered negligible. From this table, a bias of r*(S) to underestimate the correlation for correlations of .9 and .6 clearly stands out. The only other bias occurs for r*(BVT) for large sample size and correlation of .9.

Insert Table 4 about here

Table 5 contains the grand effects of the adjusted means. Generally bias is not a problem except for small sample size and a very large correlation where there is a bias to underestimate the value.

Insert Table 5 about here

In addition to MSE and bias, the efficiency of the estimators relative to the Pearson is considered,

$$\text{efficiency} = MSE(r(P)) \,/\, MSE(r^*). \qquad 4.3$$

A robust estimator should have smaller MSE than the Pearson in all contaminated situations and perform almost as well as the Pearson when no contamination is present. Table 6 contains the efficiency of the six robust estimators. When contamination is present, we can always do better than the Pearson and when the data are from a bivariate normal

distribution, the Pearson is best, as would be expected.

Insert Table 6 about here

5. DISCUSSION

The following discussion of the individual estimators draws on the information in the preceding tables.

The estimator r*(t) performed reasonably well over all situations. In more than half of the situations it has the smallest or second smallest MSE and never the largest. It is better than the Pearson in all situations except for a correlation of .9 and a contaminating standard deviation of 2, but in this case there is not much difference between the MSE for the two estimators. It also performed well in the non-contaminated situations.

The performance of r*(S) is related to the value of the population correlation. It has the smallest MSE for correlations of .0 and .3 and is almost twice as efficient as the Pearson in these situations. Its performance for correlation of .9 is worse than the Pearson. In spite of this, it never has the largest MSE. The relatively poor performance for the larger correlations is due to the bias of underestimating as mentioned above. These results are consistent with those of Wainer and Thissen, which indicated that this is the best estimator for small values of correlation but not large.

The estimator r*(BSD) is never among the best estimators and for many situations has the largest MSE. In Table 6 we can see that it is often worse than the Pearson, especially for large correlations.

Overall, r*(SSD) performed about as well or better than the Pearson. It is neither the best nor the worst estimator in any situation.

Inconsistent results occur for r*(BVT). It usually has the largest MSE, especially for correlations of .9 when its effect is approximately twice that of the grand effects for the subtables in Table 2. For correlations of .9 and .6, the MSE is larger than that of the Pearson. For correlations of .3 and .0 and a contaminating standard deviation of three, r*(BVT) has the second smallest MSE for both sample sizes and is 50% more efficient than the Pearson. For a standard deviation of two, it performs reasonably well for the small samples but not for the large. It does not perform well for the bivariate normal distributions.

The averaged bisquare, r*(B'), is usually the best estimator for correlations of .9 and .6 and is among the better estimators for the smaller correlations. In terms of efficiency, it is always better than the Pearson for the contaminated situations and is close to the Pearson in the non-contaminated situations.

6. CONCLUSIONS

In this study, the performance of many of the estimators is dependent on the value of the population correlation, sample size and/or type of contamination. The estimation of very large correlations appears to represent a special situation. Two estimators that stand out as not being situation specific and demonstrate robust properties are the averaged bisquare, r*(B') and Tukey, r*(t), correlations. The transformed Spearman's rho is an attractive estimator because of the ease of calculation; its performance however is dependent on the correlation. For small correlations it is clearly the best of all the estimators considered in this study but degenerates as the correlation increases such that it is among the worst for correlation of .9. Of the other three estimators, r*(SSD) demonstrated robustness and, given its performance in other studies, warrants further investigation. The r*(BSD) did not demonstrate robustness in this study and r*(BVT) demonstrated limited usefulness.

The results of this study are not completely comparable to some of the previous research since only one type of non-Gaussian situation is considered - contaminated normal distributions. Devlin et al. used four other types of distributions in addition to the contaminated normals. Two of the leading estimators from their study, r*(BVT), in a different form, and r*(SSD) did not perform as well as some

of the other robust estimators included in this study. The performance of r*(B') and r*(t) with respect to other types of non-Gaussian distributions is not known.

7. REAL DATA EXAMPLES

This was a Monte Carlo study in which the data were generated to possess specific distributional properties and known parameters. It is important to relate the usefulness of this type of research to real data where the distributions and parameters are unknown and the purpose of analysis is estimation.

Insert Table 7 about here

Two data sets are presented in table 7. These are student scores on midterm and final exams in intermediate (example 1) and advanced (example 2) statistical analysis classes. Figures 1 and 2 contain the scatter plots for these data.

Insert Figure 1 about here

In figure 1, there are two observations in the lower left corner. The effect of these points is that the Pearson will tend to overestimate the relationship between the variables. Table 8 contains the correlation estimates. All six of the robust estimates are smaller than that of the Pearson, minimizing the effect of the two "outliers". The

most extreme estimates are r*(BSD) and r*(BVT).

Insert Table 8 about here

In figure 2, suspicious or questionable observations are on the minor axis and would cause the Pearson to underestimate the correlation. The robust estimates are larger than that of the Pearson except for r*(S) and r*(BVT). These data might be considered to have moderately high correlation (.7) and we saw above that r*(S) has a tendency to underestimate in this situation. The values for r*(B') and r*(t), .703 and .705 respectively, are very close. The value for r*(SSD) is most extreme.

Insert Figure 2 about here

The above data sets are small but represent the ability of the robust estimators to handle different types of suspicious points. The different estimators give different estimates indicating that they treat observations differently. The application of robust estimation to real data sets and problems has just begun and it needs to be extended and continued along with Monte Carlo investigation.

References

Andrews, D.F. A Robust Method for Multiple Linear Regression. Technometrics, 1974, 16, 523-31.

Andrews, D.F., Bickel, P.J., Hampel, F.R., Huber, P.J., Rogers, W.H. and Tukey, J.W. Robust Estimates of Location. Princeton, N.J.: Princeton University Press, 1972.

Baranowski, B.B. A Comparison of Regression Approaches to Robust Correlation Estimation. Paper presented to the American Educational Research Association, Boston, Mass., April, 1980.

Chou, S.P. Confidence Interval Robustness of Correlation Coefficients. Paper presented to the Psychometric Society, Iowa City, Iowa, May, 1980.

Chou, S.P. The Application of Trimming Techniques to the Robust Estimation of Correlation. Unpublished Master's Thesis, Syracuse University, 1979.

Cochran, W.G. Sampling Techniques. New York, N.Y.: John Wiley and Sons, Inc., 1963.

Denby, L. and Larson, W.A. Robust Regression Estimators Compared via Monte Carlo. Communications in Statistics - Theory and Methodology, 1977, A6, 335-62.

Devlin, S., Gnanadesikan, R. and Kettenring, J.R. Robust Estimation and Outlier Detection with Correlation Coefficients. Biometrika, 1975, 62, 531-46.

Gnanadesikan, R. and Kettenring, J.R. Robust Estimates, Residuals and Outlier Detection with Multiresponse Data. Biometrics, 1972, 28, 81-124.

Gross, A.M. Confidence Interval Robustness with Long-Tailed Symmetric Distributions. Journal of the American Statistical Association, 1976, 71, 409-16.

Gross, A.M. Confidence Intervals for Bisquare Regression Estimates. Journal of the American Statistical Association, 1977, 72, 341-54.

Hampel, F.R. A General Qualitative Definition of Robustness. Annals of Mathematical Statistics, 1971, 42, 1887-96.

Hogg, R.V. Some Observations on Robust Estimation. Journal of the American Statistical Association, 1967, 62, 1179-86.

Huber, P.J. Robust Statistics: A Review. Annals of Mathematical Statistics, 1972, 43, 1041-67.

Huber, P.J. Robust Statistical Procedures. Philadelphia, Pa: Society for Industrial and Applied Mathematics, 1977b.

Knuth, D.E. The Art of Computer Programming Vol.2/ Seminumerical Algorithms. Reading, Mass: Addison-Wesley Publishing C., 1969.

Kraemer, H.C. On Estimation and Hypothesis Testing Problems for Correlation Coefficients. Psychometrika, 1975, 40, 473-85.

Maronna, R.A. Robust M-Estimators of Multivariate Location and Scatter. Annals of Statistics, 1976, 4, 51-67.

Mosteller, R. and Tukey, J.W. Data Analysis and Regression. Reading, Mass.: Addison-Wesley Publishing Co., 1977.

Ramsay, J.O. A Comparative Study of Several Robust Estimates of Slope, Intercept and Scale in Linear Regression. Journal of the American Statistical Association, 1977, 72, 608-15.

Stigler, S.M. Simon Newcomb, Percy Daviell and the History of Robust Estimation, 1885-1920. Journal of the American Statistical Association, 1973, 68, 872-79.

Tukey, J.W. Exploratory Data Analysis. Reading, Mass.: Addison-Wesley Publishing Co., 1977.

Wainer, H. Robust Statistics: A Survey and Some Prescriptions. Journal of Educational Statistics, 1976, 1, 285-312.

Wainer, H. and Thissen, D. Three Steps Towards Robust Regression. Psychometrika, 1976, 41, 9-34.

TABLE 1

Number of Replications

	n = 20		n = 40	
	nct	ct	nct	ct
rho				
.9	100	100	100	100
.6	300	300	100	100
.3	500	600	200	300
.0	500	600	200	300

nct - no contamination
ct - contamination

TABLE 2

MSE for r*( ) times 10000

	n = 20						n = 40						
sigma	2			3			2			3			
beta		.10	.25		.10	.25		.10	.25		.10	.25	
rho	*			*			*			*			
.9	34	-4	4	52	-9	8	14	-1	1	22	-4	4	**
r*(B')	-4	0	1	-12	1	0	-2	1	0	-5	1	-1	
r*(BSD)	-1	0	0	7	-2	3	1	0	0	1	-3	3	
r*(S)	28	1	-1	35	3	-2	9	0	0	10	-1	0	
r*(t)	-2	0	1	-6	-1	1	0	1	-1	-3	4	-4	
r*(BVT)	64	-3	4	95	-6	6	39	-1	1	48	-8	7	
r*(SSD)	0	1	-2	-13	3	-2	0	0	1	-1	0	0	
.6	371	-10	10	442	-28	28	152	-10	10	192	-22	21	**
r*(B')	-34	7	-7	-16	13	-14	-26	15	-16	-38	22	-21	
r*(BSD)	7	-8	9	64	-61	60	-2	0	1	32	-22	22	
r*(S)	-7	-1	1	-40	4	-3	2	1	-1	-4	0	0	
r*(t)	-35	2	-2	-8	10	-10	-24	12	-13	-29	25	-25	
r*(BVT)	44	1	-1	23	-3	3	38	-2	2	39	-17	18	
r*(SSD)	7	-3	2	8	-11	11	2	0	0	4	0	0	
.3	543	-14	14	659	-47	46	250	-9	9	302	-27	28	**
r*(B')	-3	-11	12	11	-16	16	1	-8	8	28	-13	14	
r*(BSD)	27	-1	2	39	-9	9	-1	-1	2	13	-4	5	
r*(S)	-35	4	-3	-106	24	-23	-23	2	-1	-59	13	-12	
r*(t)	-18	-5	5	13	9	-9	-9	-5	5	12	1	0	
r*(BVT)	3	7	-6	-58	23	-23	23	9	-10	-21	24	-24	
r*(SSD)	9	0	-1	-11	-11	12	6	3	-3	-11	0	0	
.0	644	-5	5	789	-29	30	304	-17	16	372	-36	36	**
r*(B')	9	0	1	29	-13	13	-10	-2	3	20	-8	7	
r*(BSD)	55	-4	4	31	-23	22	10	-8	8	23	-18	18	
r*(S)	-38	2	-2	-136	13	-14	-22	1	-1	-68	8	-8	
r*(t)	-9	1	0	8	24	-25	-14	-1	1	2	11	-10	
r*(BVT)	-13	1	0	-126	16	-16	32	6	-6	-36	16	-15	
r*(SSD)	27	-5	5	-8	-28	28	16	3	-3	-3	-7	7	

* estimator effect

** contamination effect

TABLE 3

Grand Effects of MSE (times 10000)

	n = 20		n = 40	
sigma	2	3	2	3
rho				
.9	34	52	14	22
.6	371	442	152	192
.3	543	659	250	302
.0	644	789	304	372

TABLE 4

Adjusted Means (times 1000)

		n = 20					n = 40					
	sigma	2		3			2		3			
	beta	.10	.25	.10	.25		.10	.25	.10	.25		
	*						*					
rho												
.9	889	4	-1	0	-7		893	2	3	-2	-3	**
r(P)	1	1	-1	3	0		0	-2	1	-3	6	
r*(B')	0	0	0	0	3		0	-1	-1	2	2	
r*(BSD)	3	1	0	0	-4		-3	0	-1	3	-5	
r*(S)	-25	1	2	0	0		-16	0	1	0	0	
r*(t)	2	-1	-1	2	1		0	-2	0	0	6	
r*(BVT)	5	-1	7	1	-11		39	8	2	-1	-26	
r*(SSD)	-3	-3	0	0	4		-3	0	-1	3	0	
.6	578	5	0	0	-10		598	1	2	-2	-1	**
r(P)	4	-1	1	0	4		3	-3	-2	1	3	
r*(B')	10	-1	0	-1	2		3	-1	0	-2	5	
r*(BSD)	-3	3	0	0	-6		-10	7	0	0	-10	
r*(S)	-16	0	1	-2	0		-18	4	-3	3	-14	
r*(t)	7	-3	0	0	5		0	0	-3	4	0	
r*(BVT)	0	4	3	-4	-6		5	9	1	-1	-17	
r*(SSD)	-4	1	-1	2	-3		-12	0	0	-2	1	
.3	280	2	0	-1	-6		295	1	2	0	-1	**
r(P)	-6	-1	3	-9	0		3	0	-1	-2	2	
r*(B')	1	0	1	-2	2		2	-1	0	-4	2	
r*(BSD)	5	-1	0	3	0		-4	1	0	0	-5	
r*(S)	-7	0	0	0	-2		-10	5	-1	1	-4	
r*(t)	0	-1	1	-5	0		2	0	0	-2	2	
r*(BVT)	-3	0	-1	0	-6		0	4	0	1	-9	
r*(SSD)	1	0	-1	1	-1		-2	0	-1	4	0	
.0	-3	0	-1	1	0		2	4	-4	3	-6	**
r(P)	3	-1	0	-1	1		0	2	-3	9	-3	
r*(B')	3	1	-2	5	-1		0	0	0	-1	-1	
r*(BSD)	0	5	1	-1	-3		-2	0	0	-1	1	
r*(S)	-5	-1	-1	0	0		-1	0	-1	1	0	
r*(t)	-1	0	1	0	0		-1	1	-2	5	-2	
r*(BVT)	-5	0	1	-1	2		3	-2	1	0	1	
r*(SSD)	0	2	-2	2	-4		1	-1	4	-2	0	

* estimator effect

** contamination effect

TABLE 5

Grand Effects of Adjusted Means

rho	n = 20	n = 40
.9	.889	.893
.6	.578	.598
.3	.280	.295
.0	-.003	.002

TABLE 6

MSE Efficiency Relative to the Pearson

	n = 20					n = 40				
sigma	2		3		1	2		3		1
beta	.10	.25	.10	.25	0	.10	.25	.10	.25	0
rho = .9										
r*(B')	1.00	1.17	1.26	1.52	.95	1.22	.98	1.98	1.08	1.00
r*(BSD)	.90	1.12	.85	.97	.88	1.07	.84	1.69	.73	1.00
r*(S)	.44	.64	.50	.73	.55	.68	.56	1.03	.60	.76
r*(t)	.93	1.11	1.12	1.24	.91	1.09	.98	1.47	1.12	1.00
r*(BVT)	.29	.39	.31	.42	.27	.29	.25	.48	.27	.23
r*(SSD)	.86	1.16	1.22	1.53	.75	1.09	.88	1.65	.86	.81
rho = .6										
r*(B')	1.07	1.10	1.30	1.28	.96	1.12	1.10	1.58	1.24	.97
r*(BSD)	.99	.95	1.28	.95	.96	1.06	.82	1.36	.71	.87
r*(S)	1.01	1.00	1.41	1.32	.81	1.02	.81	1.46	.90	.75
r*(t)	1.09	1.09	1.29	1.25	.99	1.14	1.06	1.47	1.20	.96
r*(BVT)	.88	.89	1.23	1.13	.70	.83	.65	1.26	.70	.68
r*(SSD)	.98	.96	1.30	1.15	.83	1.03	.80	1.40	.88	.81
rho = .3										
r*(B')	1.09	1.08	1.42	1.25	.96	1.14	1.13	1.55	1.27	.97
r*(BSD)	1.02	1.05	1.34	1.21	.88	1.12	1.17	1.58	1.36	.92
r*(S)	1.13	1.18	1.63	1.59	.86	1.22	1.29	1.96	1.82	.93
r*(t)	1.11	1.13	1.36	1.29	.96	1.18	1.19	1.56	1.38	.98
r*(BVT)	1.04	1.11	1.49	1.46	.77	.98	1.12	1.62	1.65	.80
r*(SSD)	1.05	1.08	1.47	1.30	.77	1.07	1.16	1.71	1.48	.82
rho = .0										
r*(B')	1.07	1.07	1.33	1.14	.93	1.16	1.10	1.56	1.18	1.00
r*(BSD)	1.00	.99	1.34	1.13	.89	1.10	1.02	1.60	1.15	.99
r*(S)	1.14	1.15	1.62	1.47	.93	1.20	1.16	1.97	1.55	.97
r*(t)	1.09	1.10	1.30	1.23	.98	1.17	1.12	1.56	1.29	1.03
r*(BVT)	1.10	1.10	1.59	1.46	.88	.98	.99	1.72	1.44	.81
r*(SSD)	1.04	1.03	1.43	1.17	.83	1.04	1.04	1.67	1.25	.90

TABLE 7

Real Data Examples

	DATA 1		DATA 2	
obs.	X	Y	X	Y
1	51	44	42	40
2	71	49	38	38
3	78	57	36	42
4	34	11	40	42
5	80	55	35	6
6	73	52	45	39
7	56	58	13	26
8	64	54	17	8
9	73	50	40	42
10	66	53	37	45
11	52	51	14	5
12	71	51	38	34
13	54	41	31	34
14	57	59	25	5
15	54	44	32	38
16	79	47	24	0
17	37	44	38	22
18	44	58		
19	61	43		
20	66	44		
21	53	48		
22	58	55		
23	54	50		
24	51	42		
25	80	59		
26	33	13		
27	80	58		
28	53	43		
29	67	52		

TABLE 8

Correlation Estimates for Real Data

	DATA 1	DATA 2
r(P)	.638	.683
r*(B')	.468	.703
r*(BSD)	.396	.710
r*(S)	.543	.667
r*(t)	.585	.705
r*(BVT)	.352	.664
r*(SSD)	.621	.812

FIGURE 1

EXAMPLE 1

PLOT OF Y*X SYMBOL USED IS *

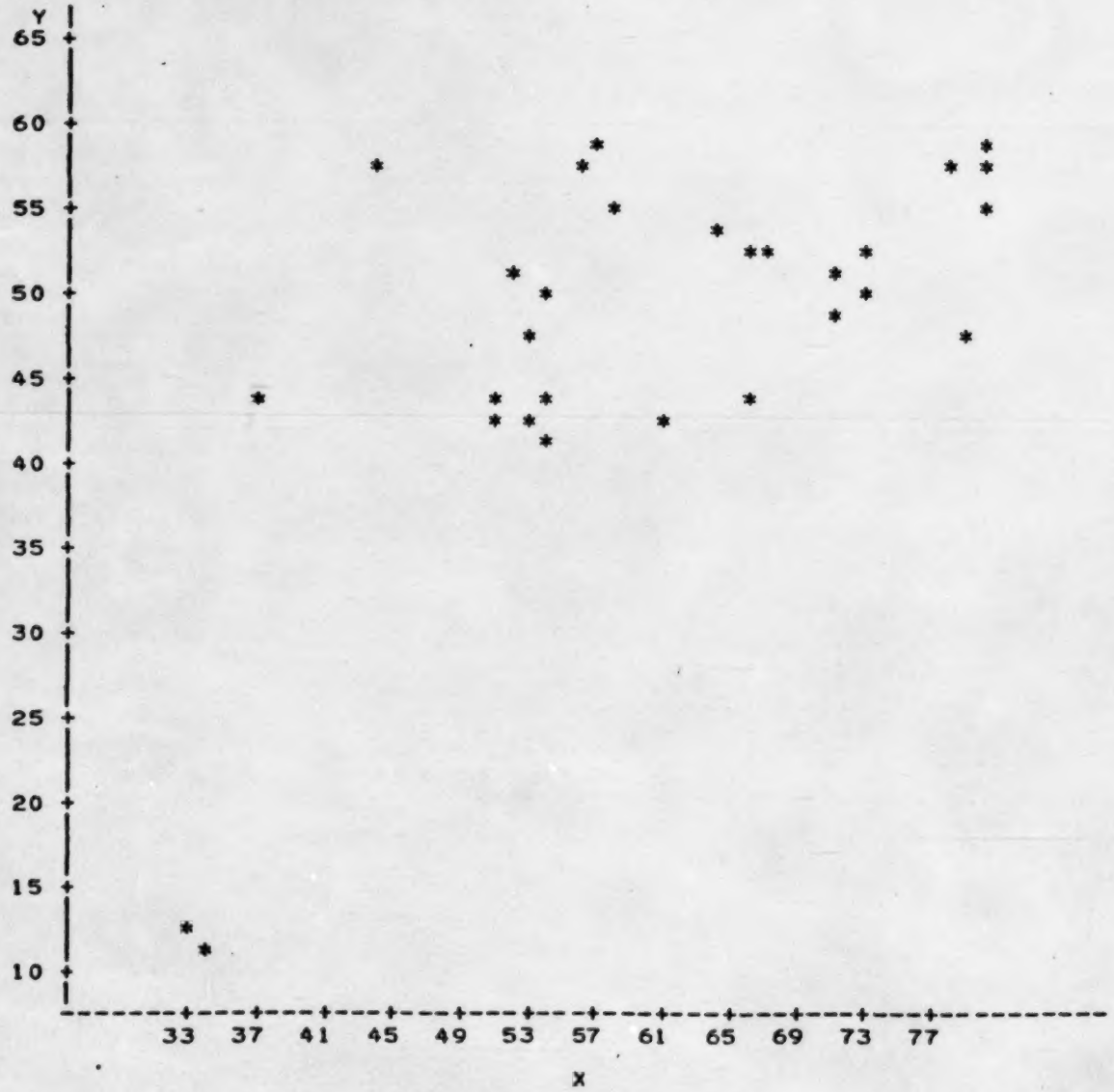
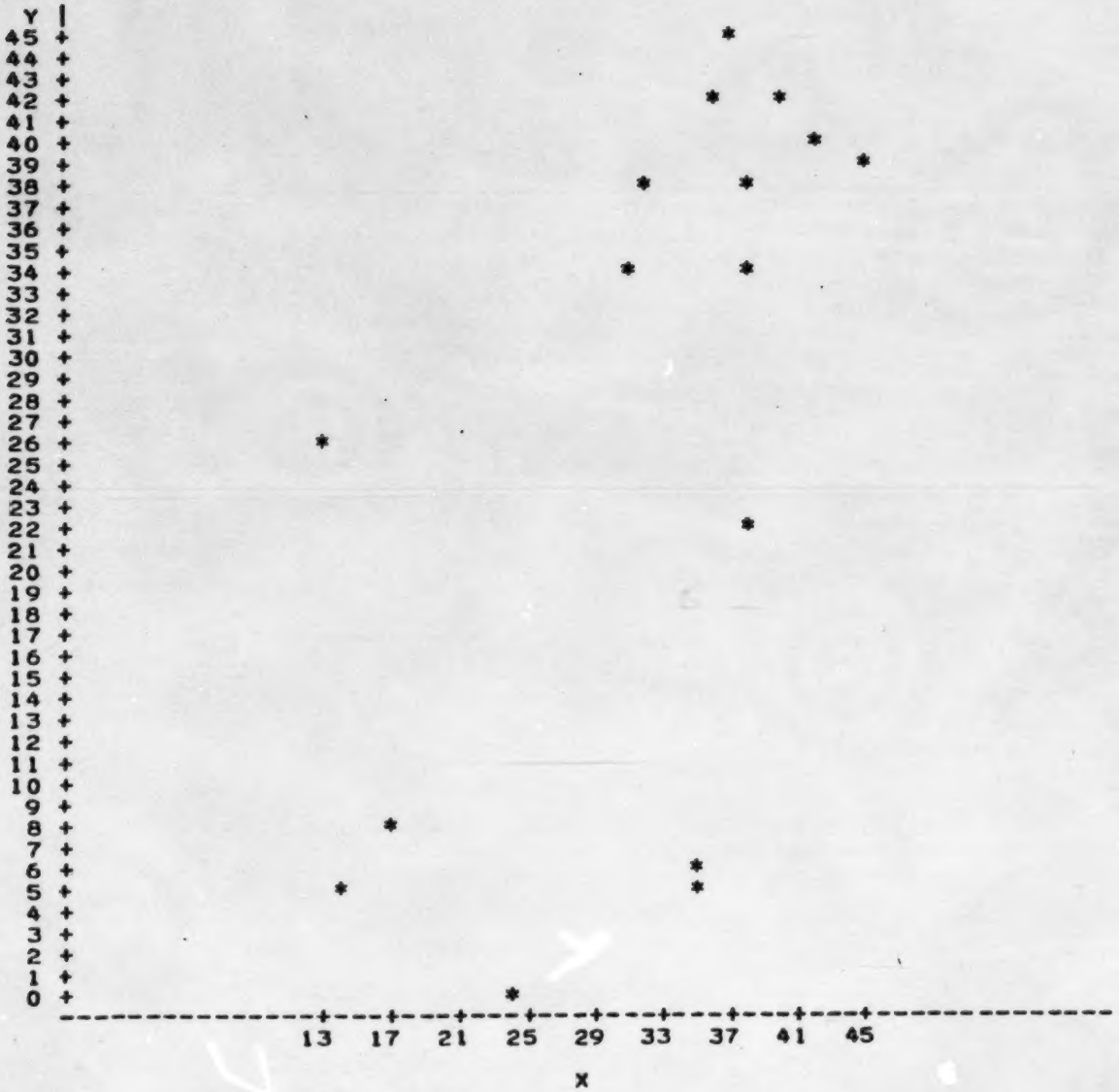


FIGURE 2

EXAMPLE 2

PLOT OF Y*X SYMBOL USED IS *



NOTE: 1 OBS HIDDEN